DOCUMENT RESUME

ED 337 537 UD 028 297

TITLE Sustained Effects of Chapter 1 Participation 1987-1990. Evaluation Report.

INSTITUTION Saginaw Public Schools, Mich. Dept. of Evaluation Services.

PUB DATE Jun 91
NOTE 33p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Achievement Gains; Compensatory Education; *Educationally Disadvantaged; Elementary Education; Longitudinal Studies; *Mathematics Achievement; *Outcomes of Education; Program Effectiveness; Program Evaluation; *Reading Achievement; Student Evaluation; Urban Schools
IDENTIFIERS *Education Consolidation Improvement Act Chapter 1; *Saginaw City School System MI; Sustaining Effects Study

ABSTRACT

This study is designed to meet the evaluation requirement of Chapter 1 of the Education Consolidation and Improvement Act of 1981, which mandated that every 3 years a determination must be made of whether or not improved performance is sustained over more than 1 year from programs funded under this Act. The Saginaw program funded by Chapter 1 is titled Academic Achievement. Its purpose is to improve the reading and mathematics achievement of a designated number of educationally disadvantaged students. Since 1979, the Academic Achievement program has served approximately 2,000 students in kindergarten through grade 9 each year. Chapter 1 activities included pretests and posttests to measure program effects, with a third data collection at least 1 year later using the California Achievement Test. Results indicate that for both reading and mathematics during the year, students generally shows gains in program participation and losses in terms of sustained effects. Recommendations are made to improve the long-term implementation and impact of the Academic Achievement program and to produce better sustained-effects results. Statistical data are presented in ten tables and four graphs. A six-item list of references is included. An appendix presents a table reflecting error in estimating achievement gains for Chapter 1 students. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED337537

EVALUATION REPORT

SUSTAINED EFFECTS OF CHAPTER 1 PARTICIPATION

1987-1990

DEPARTMENT OF EVALUATION SERVICES

- PROVIDING ASSESSMENT, PROGRAM EVALUATION AND RESEARCH SERVICES -

SCHOOL DISTRICT OF THE CITY OF SAGINAW

Saginaw, Michigan

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☒ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

RN Clams

SD Saginaw

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

UD 028 297

BEST COPY AVAILABLE

SUSTAINED EFFECTS OF CHAPTER 1 PARTICIPATION

1987-1990

An Approved Report of the

DIVISION OF ADMINISTRATION AND PERSONNEL

Department of Evaluation, Testing and Research

Richard N. Claus, Ph.D.
Manager, Program Evaluation

Barry E. Quimper, Director
Evaluation, Testing & Research

Dr. Foster B. Gibbs, Superintendent
School District of the City of Saginaw

June, 1991

TABLE OF CONTENTS

	Page
INTRODUCTION	1
STUDY DESIGN	4
PRESENTATION OF DATA	6
Total Group - Reading and Mathematics Results	7
By Gender - Reading Results	10
By Racial/Ethnic Group - Reading Results	11
By Grade Group - Reading Results	12
By Birth Year Groups - Reading Results	13
By Gender - Mathematics Results	15
By Racial/Ethnic Group - Mathematics Results	16
By Grade Group - Mathematics Results	17
By Birth Year Group - Mathematics Results	18
SUMMARY	20
RECOMMENDATIONS	23
BIBLIOGRAPHY	25
APPENDICES	26
Appendix A: Table A.1. "Give-or-Take" Table	27

LIST OF TABLES

Table		Page
1	Average Gain/Loss in Normal Equivalent Units--Reading and Math for Total Group	7
2	Average Gain/Loss in Normal Equivalent Units--Reading Results by Gender and Total Group	10
3	Average Gain/Loss in Normal Equivalent Units--Reading Results by Racial/Ethnic Group and Total Group	11
4	Average Gain/Loss in Normal Equivalent Units--Reading Results by Grade and Total Group	13
5	Average Gain/Loss in Normal Equivalent Units--Reading Results by Birth Year and Total Group	14
6	Average Gain/Loss in Normal Equivalent Units--Mathematics Results by Gender and Total Group	15
7	Average Gain/Loss in Normal Equivalent Units--Mathematics Results by Racial/Ethnic Group and Total Group	16
8	Average Gain/Loss in Normal Equivalent Units--Mathematics Results by Grade and Total Group	17
9	Average Gain/Loss in Normal Equivalent Units--Mathematics Results by Birth Year and Total Group	18
A.1	"Give-or-Take" Table	27

LIST OF FIGURES

Figure		Page
1	General Sustained Effects Results	4
2	Number Line of Reading Average Differences Pre to-Post and Post to Sustained Effects Data Points and Their Associated Confidence Intervals (N=446)	9
3	Number Line of Reading Average Differences Post to Sustained Effects Data Points and Confidence Interval for American Indians (N=2)	10
4	General Overall Pattern of Achievement During Participation and Sustained Effects Periods	22

INTRODUCTION

This study was designed to meet the evaluation requirement of Chapter 1 of the Education Consolidation and Improvement Act (ECIA) of 1981 which mandates that once every three years (after the 1978-79 school year) a determination be made whether improved performance is sustained over a period of more than one year from programs funded under this act.[1] This report represents the third in a series of reports that explores the sustaining effect of Chapter 1 programming in the School District of the City of Saginaw.

Before covering the precise details of the study an overview of Saginaw's Chapter 1 Compensatory Education Program is necessary to put the details of the study in the proper perspective.

The Saginaw program funded under Chapter 1 is currectly[2] entitled Academic Achievement (A^2). The purpose of this program is to improve the reading and mathematics achievement of a designated number of educationally disadvantaged students. The Chapter 1 funded A^2 program served approximately 2,000+ students in grades K-9 each of the eleven years since 1979.[3]

---

[1]The funding legislation was originally known as Title I of the Elementary and Secondary Education Act of 1965. Since the 1981 legislation this act was amended by the Augustus F. Hawkins - Robert T. Stafford Elementary and Secondary School Improvement Amendments of 1988 Pub. L. 100-297. However, the sustaining effect study still remains as part of the continuing legislation under this program.

[2]The 1985-86 school year was the first year the compensatory education program was evaluated under the name of Academic Achievement (A^2). Prior evaluation reports have referred to the same program as the Supplemental Teacher Participation (STP).

[3]Eighth grade students were served since the 1981-82 school year, ninth grade students have been served since the 1982-83 school year, and twelfth grade students have been served since the 1988-89 school year.

Both product and process evaluations of the program have been conducted for the past eleven years. The product evaluation reports describe the academic achievement of A^2 students fully and are available upon request from the Department of Evaluation, Testing and Research. Grade level achievement of A^2 students according to the performance standard are summarized for both subject areas in the chart below for the past eleven school years.

Overall, the achievement levels in terms of the attainment of the performance standard for the A^2 program have been very positive[4]. Achievement of the performance standard was the best during 1979-80 with a 100% positive attainment level. Over the course of the remaining ten years, two years have been above the 90% positive attainment level (1985-86 and 1986-87 at 94.4%), five years have been between 80% to 90% positive attainment levels (1988-89 at 89.5%, 1981-82 and 1983-84 at 88.9%, 1982-83 at 85%, and 1989-90 at 84.2%), two years have been between the 60% to 70% positive levels (1984-85 and 1987-88 at 66.7%) and one year has been between the 50 to 60% positive attainments levels (1980-81 at 56.2%). Student achievement levels for reading and mathematics seemed quite similar in terms of attaining the performance standard over the eleven-year period. One definite pattern, despite overall success, is the lack of achievements of the program at the seventh grade level.

---

[4]Mullin and Summers (1983) studied all the "major" compensatory education studies through 1982. Generally their review indicated that compensatory education programs have a positive though small effect on the achievement of disadvantaged students. Our findings locally through our first sustained effects study showed much larger positive gains (at or above "normal growth") across the majority of grade levels studied.

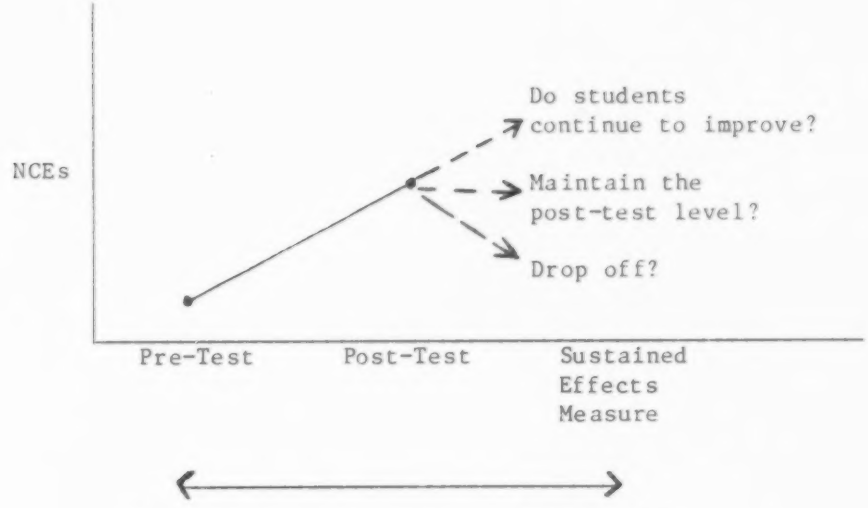
It is in this context that the Department of Evaluation, Testing, and Research has attempted to implement an investigation into the sustained effectiveness of Saginaw's Chapter 1 A^2 program on 1988-89 participants who no longer participated in the program in 1989-90. As indicated in the chart above, the 1988-89 school year participants showed a positive attainment level of 89.5% with only seventh graders (in both reading and mathematics) failing to attain the standard.

STUDY DESIGN

Generally, a sustained effects evaluation is a way to determine whether a program is effective over an extended time period. Chapter 1 evaluation activities include the administration of a pre- and post-test to determine program effects on a short-term basis. Collecting additional information at a third point in time, after the post-test, is one way to determine whether the short-term effects are sustained.

In Figure 1, the NCE gains from pre-test and post-test reflect the short-term impact of the project. By looking at a third data point, one can determine whether the project impact was sustained over a longer time period.

FIGURE 1. GENERAL SUSTAINED EFFECTS RESULTS



The purpose of any sustained effects evaluation should be to provide the Chapter 1 project with information about its impact over time, so that better decisions can be made about program planning. Some examples of programmatic

reasons for a sustained effects evaluation include questions concerning: long-term program effects; program impact by student type; early exit from Chapter 1; and impact of various types of programs.

The question addressed by this study was to determine the effect of the Academic Achievement (A^2) Chapter 1 program in reading and/or mathematics in sustaining gains for exited students in grades 1-9 over the period from Spring, 1989 to Spring, 1990 as shown by test scores on the 1985 version of the California Achievement Tests (CAT). Thus all exited 1989-90 students who participated in the reading and/or mathematics A^2 program components during the 1988-89 school year in the School District of the City of Saginaw were included in the study. CAT scores for exited students were collected Spring, 1988 (pre-test), Spring, 1989 (post-test), and Spring, 1990 (third data point). Gains in normal curve equivalent (NCE) units were calculated for the pre- to post-test and post-test to the third data point. The third data point will also be referred to as the sustained data point. The total reading and total mathematics scores of the CAT measured in normal curve equivalents (NCE) were the test scores used to gauge academic achievement gains in reading and mathematics.

Exited students test results were aggregated by the following student types: total, gender, racial/ethnic, grade, and birth year. Thus the variables of gender, racial/ethnic, grade (Spring, 1989) and birth year of student groups were used to classify exited students for additional analyses beyond the total group. These aggregations (also termed "disaggregations") by demographic variables were employed to search for information to further document sustained efforts for various groups of students.

PRESENTATION OF DATA

What follows is a presentation of the differences for pre- to post-test and post-test to sustained (third) data point CAT results of exited A^2 Chapter 1 participants. First the results of the total group in reading and mathematics will be presented. Then the reading results for exited students aggregated by gender, racial/ethnic, grade, and birth year groups will be presented and discussed. A similar presentation of the mathematics results will follow for exited Chapter 1 students.

Before getting into the results, the reader needs to be reminded that any achievement test score has some error attached to it through measurement. The extent of this error can be estimated by various formulas. The concept of error is even more important when we are speaking of the difference between two scores as in this study. To estimate the confidence interval around an observed difference score (gain or loss) in which the true score will fall 66.7% of the time a "give or take" table is provided in Appendix A. A perusal of Table A.1. in Appendix A quickly communicates to the reader that error in NCE units decreases quickly as the number of students increases. The tables in the text that follow will incorporate this confidence interval concept. An asterisk will be used to indicate a difference (gain or loss) that is larger than the 66.7% confidence interval but does not cross the zero point of the difference scale.

What follows then are the tabl s for the entire group of exited students in reading and mathematics and then these same students aggregated by demographic variables first for reading and then for mathematics results. The differences greater than the 66.7% confidence interval but not crossing zero will be denoted with an asterisk.

Total Group - Reading and Mathematics Results

Table 1 below presents the difference between the total reading and total mathematics scores of the CAT in NCE units. The gain/loss for the pre- to post-test and then the post-test to the third data point are shown. The reader should be made aware that student eligibility for a reading and/or mathematics A^2 program is based on a selection score in each area. Thus not all students will attain eligibility for both programs.

TABLE 1. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS—READING AND MATH RESULTS FOR TOTAL GROUP.

Subject Area	Grades Spring, 1988	N	AVERAGE NCE SCORE				
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Gain/Loss Pre-to-Post	Gain/Loss Sustained
Reading	1-9	446	27.8	39.3	36.0	+ 11.5*	- 3.3*
Math	1-9	280	27.4	45.5	39.7	+ 18.1*	- 5.8*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

An examination of Table 1 reveals that pre-to-post gains were positive and the 66.7% confidence limit did not cross the zero point in both reading and mathematics (11.5 and 18.1 respectively). Definite gains in reading and mathematics were made by each group respectively. The sustained differences were negative in both reading and mathematics (-3.3 and -5.8 respectively) and the confidence limits did not cross zero. Indicating the entire exiting group in both reading and mathematics lost ground (failed to sustain all the gains) one year after the A^2 treatment.

For most of us the concept of a confidence interval (as explained and illustrated in Appendix A in Table A.1 in the form of "give or take" table

from the ECIA Chapter 1 Evaluation Technical Assistance Center in Indianapolis, Indiana) is not an everyday occurrence. A further illustration as shown in the chart below of Table 1 results may be instructive.

PRE- TO POST-TEST COMPARISONS

	Number Tested (N)	Pre-to-Post Gain	Give & Take Value from Table A.1	Confidence Interval
Reading	446	+ 11.5	± .5 (estimated from table)	+11.0 to +12.5
Mathematics	280	+ 18.1	± .6	+17.5 to +18.7

POST-TEST TO THIRD DATA POINT COMPARISONS

	Number Tested (N)	Pre-to-Post Gain	Give & Take Value from Table A.1	Confidence Interval
Reading	446	- 3.5	± .5 (estimated from table)	-3.0 to -4.0
Mathematics	280	- 3.3	± .6	-2.7 to -3.9

As can be seen from the chart above, the average gain/loss as well as the number tested are needed to access the appropriate "give or take" value from Table A.1 in Appendix A. This "give or take" value is added and subtracted from the average difference to obtain a confidence interval with an upper and lower limit. Thus on successive measurements of the average differences we can expect that 66.7% of the time that the confidence interval would contain the true average difference. If we go further and graph the reading average gain/loss and confidence intervals as has been done in Figure 2 some additional insights might be apparent. A number line is represented by the line in Figure 2 below.

FIGURE 2. NUMBER LINE OF READING AVERAGE DIFFERENCES PRE-TO-POST AND POST TO SUSTAINED EFFECTS DATA POINTS AND THEIR ASSOCIATED CONFIDENCE INTERVALS (N=446).

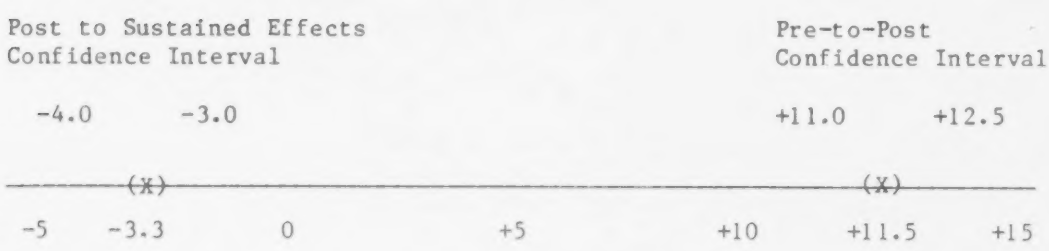


Figure 2 above illustrates the number line of average differences with the pre-to-post and post to sustained average differences and the confidence intervals around each. Since both confidence intervals do not cross the zero point, then we can have at least confidence (at the 66.7% level) that we are looking at a difference greater than zero (a gain) considering error from pre-to-post and a difference less than zero (a loss) considering error from post to sustained effects measurement. An asterisk is being used in the tables then to indicate a confidence interval that either is a clear gain (on average considering a 66.7% confidence interval students continue to improve from their initial score levels) or a clear loss (on average considering a 66.7% confidence interval students continue to drop off from their initial score level). The absence of an asterisk indicates that the confidence interval crosses the zero and the best that can be said at the stated confidence interval is that students maintain their initial score level (neither a gain or loss). Figure 3 captures a number line illustration of a confidence interval that crosses the zero point.

FIGURE 3. NUMBER LINE OF READING AVERAGE DIFFERENCES POST TO SUSTAINED EFFECTS DATA POINTS AND CONFIDENCE INTERVAL FOR AMERICAN INDIANS (N=2).

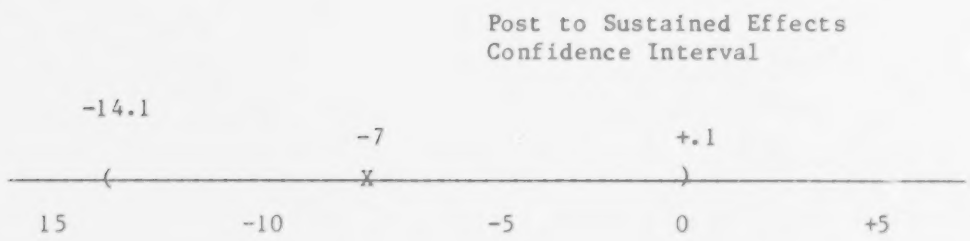


Figure 3 above illustrates an average gain that would lack an asterisk because the confidence interval crosses zero because of the error attached to the scores with a small number sampled (N=2). Thus we can say that the initial test score seems to be maintained.

By Gender - Reading Results

Table 2 below presents the reading results for exited students for males and females.

TABLE 2. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS—READING RESULTS BY GENDER AND TOTAL GROUP.

Subject Area	Gender	N	AVERAGE NCE SCORE				
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Gain/Loss	
						Pre-to-Post	Sustained
Reading	Males	228	27.7	39.4	35.6	+ 11.7*	- 3.8*
	Females	218	27.9	39.3	36.5	+ 11.4*	- 2.8*
	Total Group	446	27.8	39.3	36.0	+ 11.5*	- 3.3*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

A review of Table 2 above reveals that both males and females made approximately equal gains during the program (+11.7 and +11.4 respectively) and both males and females lost ground between the post-test and the third data point (-3.8 and -2.8 respectively). Overall, the A^2 program did not show any definite gender differences with the exception that males drop off slightly more during the sustaining period.

By Racial/Ethnic Group - Reading Results

Table 3 that follows compares reading results by racial/ethnic group for both pre-to-post and post to sustained effects data points.

TABLE 3. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS—READING RESULTS BY RACIAL/ETHNIC GROUP AND TOTAL GROUP.

Subject Area	Racial/Ethnic	N	AVERAGE NCE SCORE				
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Gain/Loss Pre-to-Post	Gain/Loss Sustained
Reading	A. Indian	2	35.5	36.5	29.5	+ 1.0	- 7.0
	Caucasian	81	39.6	44.5	40.8	+ 4.9*	- 3.7*
	Hispanic	62	27.1	40.9	39.3	+13.8*	- 1.6*
	Black	297	27.4	37.5	34.0	+10.1*	- 3.5*
	Oriental	4	28.2	52.5	37.2	+24.3*	-15.3*
	Total Group	446	27.8	39.3	36.0	+11.5*	- 3.3*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

As indicated in Table 3, the same pattern for reading is being shown. This pattern in reading shows gains for program participation from pre-to-post-test comparisons and a falling off (losses) from post to sustained effects data points. The only exception to this pattern when error in measurement is considered in Table 3 is for American Indians from pre-to-post

where all that can be said is that they appear to maintain the initial pre-test results. This inconsistent result is probably in large part due to the fact that only two American Indians were found in the exited student group. All racial/ethnic groups larger than 25 (Blacks, Caucasians, and Hispanics) appear to benefit from the A^2 treatment during the pre-to-post-test period with Hispanics benefiting the most (+13.8), followed closely by Blacks (+10.1), and then Caucasians (+4.9). During the sustaining period, all racial/ ethnic groups larger than 25 showed losses with Caucasians and Blacks showing almost equal losses of (-3.7 and -3.5 respectively) and Hispanics showing the least loss of -1.6.

By Grade Group - Reading Results

Table 4 below presents reading results by grade level group (as of Spring, 1989). The general pattern that has held in the past three tables (gains for the program participation period and losses for the sustained effects periods) basically is repeated in Table 4.

For the treatment period, a noteable exception to the above pattern is for the seventh grade group that shows a loss of -1.4, or when error is considered, shows no difference for the group of 18 students.

For the sustained effects period, a few noteable exceptions include the following: for grades 2 and 8 that show no difference when confidence interval overlap with zero is considered (+0.6 and -1.5 respectively) and for grade 7 that showed a gain of 2.6.

TABLE 4. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS—READING RESULTS BY GRADE AND TOTAL GROUP.

Subject Area	Grade Spring, 1989	N	AVERAGE NCE SCORE			Gain/Loss	
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Pre-to-Post	Sustained
Reading	1	3	21.0	59.6	42.6	+38.6*	-17.0*
	2	85	26.5	45.8	46.4	+19.3*	+ 0.6
	3	92	29.3	44.7	38.6	+15.4*	- 6.1*
	4	55	29.3	39.6	35.8	+10.3*	- 3.8*
	5	45	30.6	42.0	38.8	+11.4*	- 3.2*
	6	97	30.4	34.7	31.0	+ 4.3*	- 3.7*
	7	18	24.7	23.3	25.9	- 1.4	+ 2.6*
	8	28	20.2	29.5	28.0	+ 9.3*	- 1.5
	9	23	19.3	29.6	20.2	+10.3*	- 9.4*
Total Group		446	27.8	39.3	36.0	+11.5*	- 3.6*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

By Birth Year Groups - Reading Results

Table 5 below contrasts the reading performance of different age groups (as defined by their year of birth).

TABLE 5. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS — READING RESULTS BY BIRTH YEAR AND TOTAL GROUP.

Subject Area	Birth Year	N	AVERAGE NCE SCORE				
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Gain/Loss	
						Pre-to-Post	Sustained
Reading	1982	1	29.0	61.0	41.0	+32.0*	-19.0*
	1981	35	27.3	49.2	50.9	+76.5*	+ 1.7
	1980	72	27.8	44.8	42.7	+17.0*	- 2.1*
	1979	73	29.3	44.4	39.0	+15.1*	- 5.4*
	1978	67	28.3	40.6	37.2	+12.3*	- 3.4*
	1977	60	30.8	37.7	32.4	+ 6.9*	- 5.3*
	1976	65	30.0	34.2	31.6	+ 4.2*	- 2.6*
	1975	35	25.1	28.2	28.0	+ 3.1	- 0.2
	1974	19	18.1	28.1	25.1	+10.0*	- 3.0*
	1973	19	18.9	30.4	19.7	+11.5*	-10.7*
Total Group		446	27.8	39.2	36.0	+11.5*	- 3.3*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

An examination of Table 5 above reveals the same general pattern in the past four tables that showed gains for the program participation period and losses for the one year sustained effects period. The exceptions to this general pattern occurred for the sustained effects period. When error of the test scores is taken into account, students with the birth years of 1981 and 1975 showed a basic maintenance of the post-test levels (the actual scores were +1.7 and -0.2 respectively).

Over the birth years of 1975 through 1980 it appears that the younger the student's age the more on average each younger group gained than the proceeding older age group during the one year of program participation. A quick glance back to Table 4 (by grade) reveals a similar trend from grades 1 through 7.

What follows next are the results for mathematics for the various groups: gender, racial/ethnic, grade level, and birth year groups. At this point it should be expected that the general pattern of gains for the program participation period and losses for the one year sustained effects period will hold. Minor variations to this pattern can be expected, but the general pattern should hold as seen in Table 1 for mathematics results of the entire group with a gain of 18.1 for the program participation period and a loss of 5.8 (both differences expressed in NCE units).

By Gender - Mathematics Results

Table 6 below compares the average differences of males and females in term of total mathematics scores for the program participation period (pre- to post-test) and the sustained effects period (post-test to sustained effects data point). Thus differences reflect the time period from the spring of one year to the spring of the next year.

TABLE 6. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS — MATHEMATICS RESULTS BY GENDER AND TOTAL GROUP.

Subject Area	Gender	N	AVERAGE NCE SCORE				
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Gain/Loss	
						Pre-to-Post	Sustained
Math	Males	139	26.7	44.6	38.8	+17.9*	- 5.8*
	Females	141	28.2	46.5	40.5	+18.3*	- 6.0*
Total Group		280	27.4	45.5	39.7	+18.1*	- 5.8*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

A review of Table 6 indicates that the same general pattern that was found in the prior tables is again illustrated (gains are observed during the participation period for both males and females and losses are observed during

the sustained effects period for both males and females). During the participation period females gain more than males (+18.3 and +17.9 respectively) and females also lose more during the sustained effects period (-6.0 and -5.8 respectively).

By Racial/Ethnic Group - Mathematics Results

Table 7 presents mathematics results for the five racial/ethnic groups. As before the asterisk is used to indicate that the 66.7% confidence interval fails to cross the zero point of the difference (admittedly this is a crude way to take into account the sample size as the amount of confidence to place in an observed score(s).

TABLE 7. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS — MATHEMATICS RESULTS BY RACIAL/ETHNIC GROUP AND TOTAL GROUP.

Subject Area	Racial/ Ethnic Group	N	AVERAGE NCE SCORE				
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Gain/Loss Pre-to-Post	Gain/Loss Sustained
Math	A. Indian	1	68.0	54.0	45.0	+14.0*	- 9.0
	Caucasian	57	29.7	52.3	44.2	+22.6*	- 8.1*
	Hispanic	45	26.5	45.5	37.0	+19.0*	- 8.5*
	Black	176	26.7	43.3	38.8	+16.6*	- 4.5*
	Oriental	1	34.0	53.0	37.0	+19.0*	-16.0*
Total Group		280	27.4	45.5	39.7	+18.1*	- 5.8*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

An examination of Table 7 above reveals that again the general pattern of gains during the participation period and losses during the sustained effects period has been repeated by the racial/ethnic group. The noteable exception is for the American Indian group that during the sustained effects period had no discernable difference (-9.0, N=1) when the confidence interval is considered. Of the three racial/ethnic groups with numbers tested greater

than 25, Caucasians showed the greatest gain during participation and the second greatest loss during the sustained effects period (+22.6 and -8.1 respectively). The Hispanics showed the second greatest gain and the largest loss (+19.0 and -8.5 respectively), and the Blacks showed the third greatest gain and loss (+16.6 and -4.1 respectively).

By Grade Group - Mathematics Results

Table 8 below compares the mathematics results for the participation and sustained effects periods by grade level group (Spring, 1989 grade indicated).

TABLE 8. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS — MATHEMATICS RESULTS BY GRADE AND TOTAL GROUP.

Subject Area	Grade Spring, 1989	N	AVERAGE NCE SCORE			Gain/Loss	
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Pre-to-Post	Sustained
Math	1	3	23.6	56.6	34.6	+33.0*	-22.0*
	2	49	24.4	49.6	43.5	+25.2*	- 6.1*
	3	65	27.3	51.4	40.4	+24.1*	-16.3*
	4	60	28.9	46.6	45.9	+17.7*	- 0.7
	5	33	29.3	46.1	41.5	+16.8*	- 4.6*
	6	38	29.1	39.5	30.3	+10.4*	- 9.2*
	7	3	47.6	44.6	38.3	- 3.0	- 6.3*
	8	22	25.8	29.2	29.8	+ 3.4*	+ 0.6
	9	7	18.4	31.5	28.8	+13.1*	- 2.7
Total Group		280	27.4	45.5	39.7	+18.1*	- 5.8*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

A perusal of Table 8 reveals again the same general pattern of the prior seven tables that show gains in the participation period and losses in the sustained effects period. The exception during the participation period when considering the confidence interval occurred with grade seven that showed basically a post-test score that maintained the pre-test score level (-3.0,

N=3) and a loss for the sustained effects period (-6.3, N=3). Three noteable maintenance of gains occurred during the sustained effects period at grades nine, eight, and four (-2.7, N=7; +0.6, N=22; and -0.7, N=60 respectively). Gains during the participation period were the largest at grade one and tended to drop off as you preceded from grade two through eight (+33.0, +25.2, +24.1, +17.7, +16.8, +10.4, -3.0, and +3.4 respectively).

By Birth Year Group - Mathematics Results

Table 9 below contrasts birth year groups of 1981 through 1973 on average differences in mathematics for both the participation and the sustained effects periods. Again the general pattern holds of gains during the participation and losses for the sustained effects period.

TABLE 9. AVERAGE GAIN/LOSS IN NORMAL EQUIVALENT UNITS — MATHEMATICS RESULTS BY BIRTH YEAR AND TOTAL GROUP.

Subject Area	Birth Year	N	AVERAGE NCE SCORE				
			Pre-Test Spring, 1988	Post-Test Spring, 1989	Sustained Spring, 1990	Gain/Loss	
						Pre-to-Post	Sustained
Math	1981	25	26.1	50.6	44.5	+24.5*	- 6.1*
	1980	43	25.5	51.6	41.4	+26.1*	-10.2*
	1979	64	28.2	50.1	44.5	+21.9*	- 5.6*
	1978	49	28.4	47.4	43.8	+19.0*	- 3.6*
	1977	37	27.7	42.5	36.3	+14.8*	- 6.2*
	1976	28	31.0	40.4	31.8	+ 9.4*	- 8.6*
	1975	14	25.2	30.3	23.7	+ 5.1*	- 6.6*
	1974	14	26.3	30.6	34.8	+ 4.3*	+ 4.2*
	1973	6	20.6	31.1	27.6	+10.5*	- 3.5
Total Group		280	27.4	45.5	39.7	+18.1*	- 5.8*

*This observed gain/loss is unlikely to be due to chance because it exceeds the gain/loss which would have been predicted using the 66.7% confidence limits described earlier in the text (and illustrated in Table A.1. [Appendix A]) and does not cross the point of zero difference.

Table 9 above presents two exceptions to the general pattern seen on the prior eight tables. The first exception occurred for the sustained effects period with no real difference for 1973 birth year (-3.5) and a gain for the 1974 birth year group (+4.2). For the participation period, it appears fairly consistent for groups with 25 or over students that as age increases gains drop off (1981 = +24.5; 1980 = +26.1; 1979 = +21.9; 1978 = +19.0; 1977 = +14.8; and 1976 = +9.4).

SUMMARY

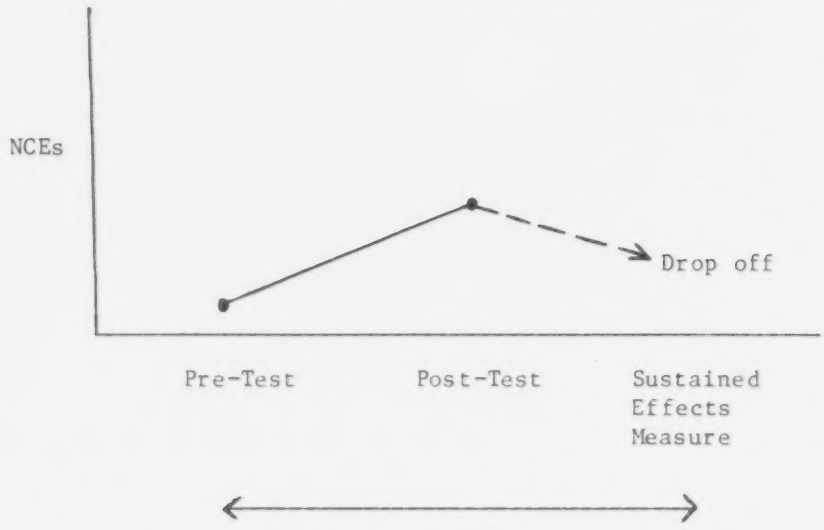
This study determined the effect of the Academic Achievement (A^2) Chapter 1 program in reading and/or mathematics in sustaining gains for exited students in grades 1-9 over the period from Spring, 1989 to Spring, 1990. Test scores on the 1985 version of the California Achievement Tests (CAT) Forms E and F were used as achievement outcome measures on Total Reading and Total Mathematics scores. Thus all exited 1989-90 students who participated in the reading and/or mathematics A^2 program components during the 1988-89 school year in the School District of the City of Saginaw were included in the study. Exited students' test results were aggregated by the following student types: total, gender, racial/ethnic, grade, and birth year. CAT scores for exited students were collected Spring, 1988 (pre-test), Spring, 1989 (post-test), and Spring, 1990 (third data point).

The chart below summarizes the participation period (1988-89) and the sustained effects period (1989-90) differences. The differences are stated considering the 66.7% confidence interval and whether this interval crosses the zero difference point. A plus (+) symbolized a gain, a minus (-) symbolizes a loss, and a zero (0) indicates no difference or maintenance situation. Both reading and mathematics subject area results are presented below.

Variable/Group	Reading		Mathematics	
	Participation	Sustained	Participation	Sustained
Gender				
Male	+	-	+	
Female	+	-	+	-
Racial/Ethnic				
American Indian	0	0	+	0
Caucasian	+	-	+	-
Hispanic	+	-	+	-
Black	+	-	+	-
Oriental	+	-	+	-
Grade				
1	+	-	+	-
2	+	0	+	-
3	+	-	+	-
4	+	-	+	0
5	+	-	+	-
6	+	-	+	-
7	0	+	0	-
8	+	0	+	0
9	+	-	+	0
Birth Year				
1982	+	-	+	-
1981	+	0	+	-
1980	+	-	+	-
1979	+	-	+	-
1978	+	-	+	-
1977	+	-	+	-
1976	+	-	+	-
1975	0	0	+	+
1974	+	-	+	0
1973	+	-	+	-
Entire Group Total	+	-	+	-

A review of the chart above reveals for both reading and mathematics that students generally gain during the year of program participation and lose during the year period of sustained effects. Figure 4 below presents this situation graphically.

FIGURE 4. GENERAL OVERALL PATTERN OF ACHIEVEMENT DURING PARTICIPATION AND SUSTAINED EFFECTS PERIODS.



The variations seen in the chart above from what is presented in Figure 4 may be due to small numbers per group ($N < 25$) with minor exceptions. Overall, the A^2 program for most exited students in reading and mathematics experiences a drop off of achievement during the sustained effects period (-3.3 and -5.8 NCEs respectively). During the participation period students in the reading and mathematics A^2 program experience improvement (+11.5 and +18.1 NCEs respectively).

RECOMMENDATIONS

Listed below are a series of recommendations based on the findings of this study. These recommendations are offered in an effort to improve the long-term implementation and impact of the Chapter 1 A^2 program plus produce better sustained effects studies.

The recommended ideas and techniques offered below stem from perceived problems and are just a few of many ways to improve the performance of the program. These recommendations represent a starting point for the change process. As solutions are sought for optimum program operations, a dialogue/discussion should be undertaken to determine the best and most workable way to solve the perceived problem. The staff, state level consultants/experts, and evaluators should be brought into these discussions so that all involved feel part of the proposed new operation of the program.

1. The results of this study strongly indicate that students should be grandfathered if they score above the selection cut score after one or more years of Chapter 1 service. The grandfathering service if properly designed, would insure that the gains would continue from Chapter 1 service to such a time when students could sustain gains. Present Chapter 1 legislation allows for a two-year grandfathering period and hopefully in the case of most students that would be an adequate length of service to help promote future positive academic achievement.

2. Detailed plans for any future sustained effects studies should be outlined in advance by the program director and staff to ensure that accurate longitudinal records of Chapter 1 participants are available. Such records would yield more comprehensive accounting of all aspects of participation (past history of participation, longitudinal test file linkage with unique student numbers for each student, determination of research/evaluation questions of interest, necessary testing points to answer questions posed, etc.) and ultimately make possible better understanding of the nature of any sustained effects through better controls.

3. The Michigan Department of Education should collect the existing sustained effects studies state-wide. A summarization of these results could then be shared with the State's educators. Also, suggestions on relevant issues and possible methodologies to use in such studies would be helpful to both small, medium, and large school districts.

4. The program director and her staff should study the aggregated, by student, characteristics to determine if further conclusions and recommendations are warranted for program operations. The Evaluation Department stands ready to help with further aggregations of the data from the sustained effects study if needed. The data from this study like so many other studies really can have a great deal more meaning and usefulness to program providers if they take the time to study the results and then act upon the implications in providing future program services.

BIBLIOGRAPHY

Burton, R. Discussion: Underlying Assumptions and Resulting Biases. In G. Echtermacht (Ed.), New Directions For Testing and Measurement: Measurement Aspects of Title I Evaluations, San Francisco: Jossey-Bass, Inc., 1980.

Echternacht, G. Title I Evaluation and Reporting System: Development of Evaluation Models. In G. Echtermacht (Ed.), New Directions for Testing and Measurement: Measurement Aspects of Title I Evaluations, San Francisco: Jossey-Bass, Inc., 1980.

Hill, P. Summer Drop-Off and the Effectiveness of Compensatory Instruction. (Paper prepared for the National Conference on Urban Education, held in St. Louis, Missouri, July 10-14, 1978), St. Louis, MO: CEMREL, Inc., 1979.

Linn, R. Discussion: Regression Toward the Mean and the Interval Between Test Administration. In G. Echtermacht (Ed.), New Directions For Testing and Measurement: Measurement Aspects of Title I Evaluations, San Francisco: Jossey-Bass, Inc., 1980.

Mullin, S. and Summers, A. Is More Better? The Effectiveness of Spending on Compensatory Education, Phi Delta Kappan, 1983, 64 (5), 339-347.

Roberts, A. Regression Toward the Mean and the Regression-Effect Bias. In G. Echtermacht (Ed.), New Directions For Testing and Measurement: Measurement Aspects of Title I Evaluations, San Francisco: Jossey-Bass, Inc., 1980.

APPENDICES

APPENDIX A

TABLE A.1. "GIVE-OR-TAKE" TABLE

Number of Students	Error (NCEs)	Number of Students	Error (NCEs)
1	10.1	25	2.0
2	7.1	30	1.8
3	5.8	35	1.7
4	5.1	40	1.6
5	4.5	45	1.5
6	4.1	50	1.4
7	3.8	55	1.4
8	3.6	60	1.3
9	3.4	65	1.3
10	3.2	70	1.2
11	3.0	75	1.2
12	2.9		
13	2.8		
14	2.7		
15	2.6	100	1.0
16	2.5	150	0.8
17	2.4	200	0.7
18	2.4	300	0.6
19	2.3		

Using the "Give-or-Take" Table for the 66.7% Confidence Interval

The Give-or-Take values in the table above reflect the error in estimating achievement gains for Chapter 1 students. The more students in the group, the more confident we can be in the accuracy of the gain.

A class of 16 Chapter 1 students had a gain of 8 NCEs last year. Look up 16 in the "Give-or-Take" table to find the value of 2.5 NCEs. Think of this as 8 NCE gain give or take 2.5 NCEs. Thus the true gain for this class is probably between 5.5 NCEs (8 -2.5) and 10.5 NCEs (8 +2.5).